

## RESOURCE ARTICLE

# For all audiences: Incorporating immature stages into standardised spider inventories has a major impact on the assessment of biodiversity patterns

Marc Domènech<sup>1</sup>  | Owen S. Wangensteen<sup>2</sup> | Alba Enguñados<sup>1</sup> |  
Jagoba Malumbres-Olarte<sup>3,4</sup> | Miquel A. Arnedo<sup>1</sup> 

<sup>1</sup>Department of Evolutionary Biology, Ecology and Environmental Sciences, Biodiversity Research Institute (IRBio), Universitat de Barcelona, Barcelona, Spain

<sup>2</sup>Norwegian College of Fishery Science, The Arctic University of Norway, University of Tromsø, Tromsø, Norway

<sup>3</sup>cE3c – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, São Pedro, Terceira, Açores, Portugal

<sup>4</sup>LiBR – Laboratory for Integrative Biodiversity Research, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

## Correspondence

Marc Domènech, Department of Evolutionary Biology, Ecology and Environmental Sciences, IRBio, Universitat de Barcelona, Barcelona, Spain.  
Email: [mdomenan@gmail.com](mailto:mdomenan@gmail.com)

## Funding information

Generalitat de Catalunya, Grant/Award Number: 2017SGR73; Organismo Autónomo de Parques Nacionales, Grant/Award Number: 485/2012; Universitat de Barcelona

**Handling Editor:** Simon Creer

## Abstract

Although arthropods are the largest component of animal diversity, they are traditionally underrepresented in biological inventories and monitoring programmes. However, no biodiversity assessment can be considered informative without including them. Arthropod immature stages are often discarded during sorting, despite frequently representing more than half of the collected individuals. To date, little effort has been devoted to characterising the impact of discarding nonadult specimens on our diversity estimates. Here, we used a metabarcoding approach to analyse spiders from oak forests in the Iberian Peninsula, to assess (1) the contribution of juvenile stages to local diversity estimates, and (2) their effect on the diversity patterns (compositional differences) across assemblages. We further investigated the ability of metabarcoding to inform on abundance. We obtained 363 and 331 species as adults and juveniles, respectively. Including the species represented only by juveniles increased the species richness of the whole sampling in 35% with respect to those identified from adults. Differences in composition between assemblages were greatly reduced when immature stages were considered, especially across latitudes, possibly due to phenological differences. Moreover, our results revealed that metabarcoding data are to a certain extent quantitative, but some sort of taxonomic conversion factor may be necessary to provide accurate informative estimates. Although our findings do not question the relevance of the information provided by adult-based inventories, they also reveal that juveniles provide a novel and relevant layer of knowledge that, especially in areas with marked seasonality, may influence our interpretations, providing more accurate information from standardised biological inventories.

## KEYWORDS

Araneae, diversity, DNA barcoding, Iberian Peninsula, metabarcoding, spiders

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Global human activity is altering the species richness and abundance of biological communities (Barlow et al., 2016; Stuart Chapin et al., 2000; Socolar et al., 2019). These disturbances are accelerating the invasion rate of exotic species (Hulme, 2009) and driving numerous species to extinction, sometimes even before they are described, in what has been referred to as the sixth mass extinction (Barnosky et al., 2011; Dirzo et al., 2014). Bioinventories and early detection methods for monitoring ecosystem changes are essential to identify and tackle unanticipated threats to biodiversity (Barnosky et al., 2012; Telfer et al., 2015). However, sampling and identifying highly abundant and diverse groups such as arthropods is a daunting task. Several methods have been devised to overcome this limitation. For instance, rapid biodiversity assessment protocols are fast yet efficient sampling strategies specifically designed to retrieve the greatest amount of information from a particular area, minimising the number and length of sampling periods (Oliver & Beattie, 1996).

Arthropods represent the largest and most abundant component of animal biodiversity. Therefore, no biodiversity monitoring programme can be considered credible unless it takes them into account (Taylor & Doran, 2001). Moreover, because of their high reproductive rates and short generation times, arthropods have the potential to inform on biodiversity changes at finer spatial and shorter temporal scales than vertebrates (Kremen et al., 1993; Yen & Butcher, 1997). However, the poor taxonomic knowledge of many arthropods limits their use as bioindicators, and they are frequently underrepresented in biodiversity assessments and conservation programmes (Cardoso et al., 2012). On the other hand, recent studies suggest that arthropod populations, which play a fundamental role in ecosystem functioning, are declining at an alarming rate (e.g., Leather, 2018).

Because of their rapid and divergent evolution, male copulatory apparatus and, to a lesser degree, female external copulatory organs are the main structures for species identification across most arthropod groups (Eberhard, 1985). Spiders are no exception. Morphological taxonomic identification is almost exclusively based on genitalic characters, that is, the structure of the copulatory bulb in males and the vulva and epigyne (external modifications of the genital area) in females. These features are only visible after the last moult, which makes immatures difficult or impossible to identify at species level (Coddington et al., 1996; Dobyns, 1997). Thus, in most inventories and diversity studies immature stages are discarded during sorting. However, immature specimens may account for between 40% and 70% of the collected specimens in biodiversity surveys (Cardoso et al., 2004; Malumbres-Olarte, Boieiro, et al., 2020; Malumbres-Olarte et al., 2019; Russell-Smith & Stork, 1995; Silva, 1996; Soukainen et al., 2020), or even up to 94% in extreme cases (Kuntner & Baxter, 1997). Disregarding immatures may significantly influence the inference of the temporal and spatial patterns of biodiversity, so their incorporation is desirable to obtain reliable estimates of diversity in short-term sampling protocols (Sørensen et al., 2002; Toti et al., 2000).

It is known that the life cycles of different species of spiders differ in the number of generations per year and in the time of the year when they are present as adults or juveniles (Aitchison, 1984; Nadal et al., 2018; Schaefer, 1976). Phenological differences can also be found even within the same species. For example, some wolf spiders are known to have annual life cycles maturing as adults in March–April and reproducing in May–June, although they can also have two clutches in the same year depending on the weather conditions (Rádai et al., 2017). Since rapid biodiversity assessment protocols are usually conducted once, they provide a “snapshot” of the population present in a certain area at a particular time. Hence considering only adult individuals would completely dismiss all the species that are present as immature stages in that particular time of the year. Few studies have partially addressed the effect of incorporating juvenile spider stages in diversity estimates. Norris (1999), in a study that included only a few species that could reliably be morphologically identified at immature stages, already pointed out that numerous species were only found as juveniles and that relative abundances changed drastically when these immature stages were considered. Pétilion et al. (2018), in a study focused on the family Ctenidae, found that significant differences in species distribution were detected only when juveniles were considered.

The use of DNA-based approaches for species identification, for example DNA barcoding (Hebert et al., 2003), ease the identification of immature stages. DNA barcode sequences of immature individuals can be assigned to species through comparison to reference databases containing barcodes of adult-based morphologically identified species (Meiklejohn et al., 2012; Richard et al., 2010). However, this technique still poses some drawbacks. For example, it requires manually extracting and amplifying each specimen individually, which for large samplings with hundreds or thousands of juveniles can be very time-consuming. Moreover, the economic cost of extracting, amplifying, and sequencing such a large amount of samples would also be considerable.

DNA metabarcoding is a more recently developed molecular technique consisting in the automated identification of multiple species from a single bulk sample containing entire or partial organisms or from environmental samples (water, soil, etc.) carrying remains of DNA (e.g., Bohmann et al., 2014; Morinière et al., 2016; Yu et al., 2012). This approach represents a clear advantage with respect to DNA barcoding, as it allows the simultaneous processing of many specimens at once, greatly reducing the workload and processing time. In addition, it is more cost efficient for large numbers of specimens, as the number of sequences obtained from a single metabarcoding run is in the order of millions (Sales et al., 2020; Watts et al., 2019). The downside of using this approach is that individual specimens cannot be traced back or are sometimes even lost in the process of preparing the bulk sample, making it impossible to revise the voucher specimens if interesting sequences were found.

Another potential drawback of the use of metabarcoding for biodiversity assessment is its presumed inability to provide abundance information. To what extent the number of sequence reads of a certain taxon correctly represents its abundance or biomass

in the sample has been a matter of much debate. Several studies have specifically addressed this issue (Deagle et al., 2019; Elbrecht & Leese, 2015; Lamb et al., 2019; Piñol et al., 2015), but the answer remains inconclusive. While some studies consider the quantitative power of metabarcoding limited (Elbrecht & Leese, 2015; Piñol et al., 2015), others found metabarcoding to give an accurate estimate of a taxon's abundance under certain conditions (Ratcliffe et al., 2020) or by applying correction factors that may vary among taxa (Kennedy et al., 2020; Thomas et al., 2016), or have even used it to quantitatively analyse dietary data (Soininen et al., 2015). One explanation for these different conclusions may be that the range of concentrations analysed varies considerably across studies, as suggested by Deagle et al. (2019). While a positive relationship between number of reads and biomass is commonly found, only a certain part of the variation in the number of reads seems to be explained by differences in the biomass in the sample, while the rest of the variation seems to be due to factors such as primer specificity (Elbrecht & Leese, 2015), different extraction success between different tissues or species (Schiebelhut et al., 2017) or the efficiency of the blocking primers of predator DNA in the case of diet studies (Piñol et al., 2015).

Here, we aimed to evaluate the impact of including juveniles of spiders in data/analyses on understanding the diversity and structure of communities. Specifically, we first quantified the additional diversity that immature spiders contribute to diversity estimates. Second, we investigated how the inclusion of immature specimens affects diversity patterns (compositional differences) across assemblages and identify the possible environmental factors responsible for such patterns. Finally, we assessed the ability of metabarcoding to recover abundance information from bulk samples by comparing the proportion of reads of a certain family with the proportion of

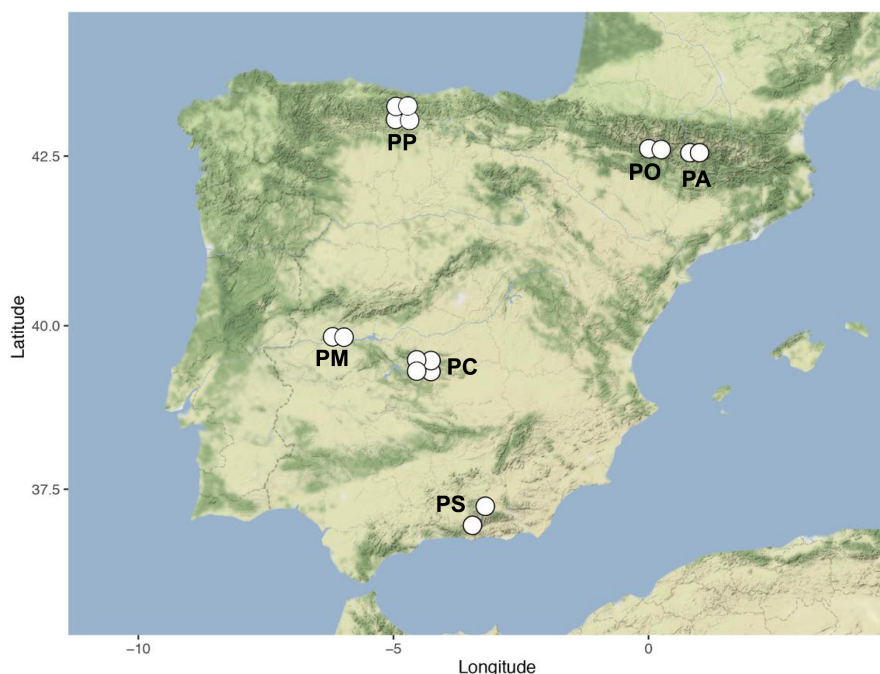
individuals and mass they represent in the sample. Our results provide important insights into the relevance of considering all different life stages in rapid biodiversity assessment protocols, which are essential to efficiently monitor ecosystem changes, and will contribute to refine the use of metabarcoding approaches as efficient alternatives to traditional, morphology-based standardised biological inventorying and monitoring schemes.

## 2 | MATERIALS AND METHODS

### 2.1 | Sample collection and sorting

We collected the specimens using the standardised sampling protocol COBRA (Cardoso, 2009) in May–June 2013 and 2014 (Crespo et al., 2018; Malumbres-Olarte, Crespo, et al., 2020). The sampling design included sixteen 1 ha plots distributed in white oak forests across six National Parks of the Iberian Peninsula, namely Aigüestortes i Estany de Sant Maurici (PA), Ordesa y Monte Perdido (PO), Picos de Europa (PP), Monfragüe (PM), Cabañeros (PC) and Sierra Nevada (PS) (Figure 1). We used semi-quantitative methods that combined 12 person-hours of timed direct capture, beating and sweeping with 48 pitfall traps active for two weeks in every plot. Detailed information on the sampling plots can be found in Table S1.

We sorted adult and immature specimens and we identified them under a ZEISS Stemi 2000 stereomicroscope. We identified adults to species level and juveniles to family level. All the immature individuals of each plot were weighed separately for each family using an analytical balance. We placed the specimens in absorbent paper for 30 min before weighing them to allow the alcohol remaining in the bodies to evaporate.



**FIGURE 1** Location of the 16 studied plots. PA, Aigüestortes i Estany de Sant Maurici; PC, Cabañeros; PM, Monfragüe; PO, Ordesa y Monte Perdido; PP, Picos de Europa; PS, Sierra Nevada

## 2.2 | DNA extraction, amplification and sequencing

Representative DNA barcode sequences for the cytochrome c oxidase subunit I (COI) (~658 bp) of species captured as adults were available from a previous study (Crespo et al., 2018) (see Table S2 and Data Availability).

For each plot, we homogenised all the collected juveniles with the help of liquid nitrogen. Two subsamples of 0.3 g were extracted from each homogenised plot using a PowerSoil DNA Isolation kit (Qiagen). We added one negative (distilled water) and one positive control, a specimen of the cobweb spider *Simitidion simile* (Koch, 1836). These controls were included in the batch, processed and sequenced along with the rest of the samples. We cleaned and sterilised all the equipment with diluted sodium hypochlorite between successive sample extractions. We amplified the COI “Leray fragment” of 313 bp using the degenerate primer set Leray-XT (Wangenstein et al., 2018). This set includes the reverse primer jgHCO2198 5'-TAIACYTCIGGRTGICCAARAAYCA-3' (Geller et al., 2013) and the forward primer mCOLintF-XT 5'-GGWACWRGWTG RACWITITAYCCYCC-3', modified from the mCOLintF primer (Leray et al., 2013). Each primer pair included twin 8-base sample tags (the same tag in the forward and reverse primers), which had at least three different base pairs (bp) between them, and a lead of 2–4 random Ns in the 5' end for increasing sequence variability of the library. The PCR mix included 10 µl AmpliTaq Gold 360 Master mix (Applied Biosystems), with 1 µl of each 5 µM forward and reverse primers, 0.16 µl of bovine serum albumin, 2 µl of DNA template and DNase-Free water to adjust the volume up to 20 µl per sample. The PCR profile included 10 min at 95°C, 35 cycles of 94°C 1 min, 45°C 1 min and 72°C 1 min, and 5 min at 72°C. We performed two PCR replicates for each extraction in the study, each with a unique tag, giving a total of four replicates per plot. For plots PA1 and PS1 we obtained three replicates (only one PCR replicate was performed for one of the two subsamples of the plot) as they were the first to be processed and served as a test. We evaluated the quality of amplifications by electrophoresis in 1% agarose in Tris-borate-EDTA buffer and stained with GelRed nucleic acid gel stain (Biotium). We pooled all PCR products by equal volume (including two PCR-negative controls and one PCR-positive control) and purified the pool using a MinElute PCR purification kit (Qiagen). Three µg of DNA from the purified pool (determined by Qubit fluorometric quantitation dsDNA BR assay kit, Thermo Fisher Scientific) were used to build a library using the NextFlex PCR-free DNA-seq kit (Perkin-Elmer). The multiplexed library was sequenced on an Illumina MiSeq with a V3 2 x 250 bp paired-end partial run at the University of Salford, UK.

## 2.3 | Bioinformatic analysis

We conducted the bioinformatic analyses using the OBITOOLS METABARCODING package (Boyer et al., 2016). We aligned the paired-end reads using the command `illuminapairedend`. We selected

sequences with alignment quality scores bigger than 40 and we demultiplexed the aligned data set and removed the primer sequences with `ngsfilter`. We also filtered out sequences containing ambiguous bases. We then used `Obiuniqu` to dereplicate the reads (grouping all identical sequences) while keeping track of their abundances, and we also removed chimeric sequences using the `uchime_denovo` algorithm in `VSEARCH` (Rognes et al., 2016). We used the step-by-step aggregation clustering algorithm implemented in `SWARM` 2.1.13 (Mahé et al., 2015) to cluster the sequences into molecular operational taxonomic units (MOTUs). For making adult (morphology and DNA barcode data) and juvenile (metabarcoding data) clustering comparable, we combined the sequences from both life stages before running the `Swarm` clustering algorithm. In the case of adults, we kept only the segment of the original COI sequences matching the Leray COI fragment. To prevent the program from discarding adult sequences as singletons, we artificially increased their initial abundance to 50,000 reads. We set a distance value of  $d = 13$  for the clustering algorithm, which has been shown to be the optimal value for discriminating intra- and interspecific divergences, that is, to approximate MOTUs to species-level clusters, in a wide range of eukaryotic systems (Antich et al., 2021; Garcés-Pastor et al., 2019; Kemp et al., 2019; Siegenthaler et al., 2019; Wangenstein et al., 2018). The species present as adults whose sequences were clustered together by `Swarm` (nine pairs, one triad and one tetrad) were also treated as single entities in downstream analyses with juveniles. One of the species of the tetrad (*Xysticus nubilus*) was only found as juvenile and not as adult. After removing the singletons, we performed the taxonomic assignment of the representative sequences of each MOTU (seeds) using `ECOTAG` (Boyer et al., 2016). We built the local reference sequence database required by `Ecotag`, combining our sequences of adult spiders with sequences retrieved from the BOLD database (Ratnasingham & Hebert, 2007) and the EMBL repository (Kulikova et al., 2004). `ECOTAG` (Boyer et al., 2016) uses a phylogenetic assignment protocol, based on the NCBI taxonomy tree, to assign sequences to the last common ancestor of the most closely related sequences in the local reference database. This approach does not require establishing arbitrary identity thresholds for every taxonomic rank (Bakker et al., 2019).

We filtered out putative contaminants of the resulting database by retaining only the MOTUs assigned to the order Araneae. After the taxonomic assignment made by `Ecotag`, we manually checked if there were better, more recent matches in BOLD or NCBI, and we updated the identification of those MOTUs for which better matches were found. We discarded as contaminants 16 MOTUs with low numbers of reads that corresponded to a checklist of non-Iberian species that had been analysed in other studies conducted in the same laboratory. At this stage we also removed the adult sequences added artificially, which eliminated 155 MOTUs. We used the `LULU` algorithm (Frøslev et al., 2017) to remove the MOTUs corresponding to pseudogenes. We also built a COI tree using the seed sequence of every MOTU and the COI sequence of the adult specimens to help allocate unassigned

MOTUs to specific families, genera or species. We inferred the tree by Maximum Likelihood using IQ-TREE v.1.6 (Nguyen et al., 2015). We partitioned positions by codon and assigned an unlinked GTR model to each partition, and we assessed branch support by means of 1000 ultrafast bootstrap approximation replicates (Hoang et al., 2018; Minh et al., 2013). Analyses were run remotely at the CIPRES Science Gateway (Miller et al., 2010). All the replicates of each plot were added up. All the MOTUs with less than five total reads were discarded. Also, for a MOTU to be counted as present in a plot, we required at least five reads in the plot and detection of the MOTU in at least two of the replicates of the plot.

## 2.4 | Delimitation of adult and juvenile clusters

For the MOTUs that could not be identified to species level, we used the best match to which the taxonomic assignment algorithm assigned that MOTU, that is, the identifier of the specific sequence in the database which was the most similar to the seed sequence of the MOTU. If the best match of two unidentified MOTUs was the same, we collapsed them and treated them as the same taxon. For comparative purposes, we also analysed the results in two additional alternative ways: using a "splitter" approach (every unidentified MOTU as a different species) and using a "lumper" approach (considering as the same species all the unidentified MOTUs of the same genus in one case, and of the same family in the other). However, using either of those approaches only translated into minor differences in the results with respect to the "best match" approach.

## 2.5 | Evaluation of the effect of including juveniles on community patterns

After processing the reads, we checked the completeness of the sequenced replicates by means of rarefaction curves, plotting the number of MOTUs per replicate against an increasing number of reads (Figure S1). To find out if there were significant differences in the similarity patterns among assemblages when including juveniles, we performed a nonmetric multidimensional scaling (NMDS) analysis based on the assemblage composition. We performed these analyses with presence/absence data both with the information on adults alone and with adults and juveniles together. For this, we used the metaMDS function in the package "VEGAN" (Oksanen et al., 2010) in R (R Core Team, 2020). We applied a Mantel test to assess the correlation and significance between the distance matrices obtained from both approaches. We also applied an analysis of similarity (ANOSIM) analysis to test if differences in species composition between northern (Aigüestortes, Ordesa and Picos de Europa) and southern parks (Monfragüe, Cabañeros and Sierra Nevada) were equally recovered by adults and adults+juveniles approaches.

## 2.6 | Assessment of species abundance from metabarcoding data

To assess the level of variation in the number of reads explained by the proportion of a certain taxa in the sample, we calculated the proportion of reads or relative read abundance (RRA), the proportion of biomass and the proportion of individuals of every family in every plot. We used the betareg function in the R package "BETAREG v1.1" (Cribari-Neto & Zeileis, 2010) to apply beta regression models to each family present in at least 10 plots. Two models were applied to each family, one for the RRA as a function of the percentage of mass the family represents in the sample, and one for the RRA as a function of the percentage of abundance. We applied the models to each family separately due to the nonindependence of percentages in a sample, and we used 10 as the minimum presence in plots following the one-in-ten rule tested in other models (Peduzzi et al., 1996). Beta regression models are designed for response variables with proportional data between 0 and 1. We calculated the adjusted  $R^2$  values and determined the global goodness of fit for each model. Only significant ( $p$ -value  $< .05$ ) models with a pseudo- $R^2 > .5$  and randomly distributed residuals were considered (Yellareddygarri et al., 2016). For each family, we calculated the average mass of a juvenile individual by dividing the total weight of the family by its number of specimens, to assess if it had an effect on the differences in reads recovered across families.

## 3 | RESULTS

### 3.1 | Sequencing results

We retrieved DNA barcodes of 368 out of the 376 species represented as adults in the sampling plots—eight species could not be sequenced. After the cleaning, filtering, and chimaera removal process, the sequencing of juvenile pool samples generated a total of 15,805,993 sequence reads and 3,839,513 unique sequences. After manually adding the sequences of the adults, the Swarm algorithm grouped all sequences in 4668 nonsingleton MOTUs, of which 140 contained exclusively adult artificial sequences. After the curation process—retaining only Araneae, removing non-Iberian taxa, and removing artificially added adult sequences—the final data set for juveniles consisted of 9,956,432 reads distributed in 1343 MOTUs. The process of removing pseudogenes eliminated 455 MOTUs and left 888 final juvenile MOTUs (final data set in Table S3). Sequencing of three of the four replicates of plot PM1—the two PCR replicates of one of the subsamples and one PCR replicate of the other subsample—was unsuccessful due to unknown reasons, so we excluded them from downstream analyses.

The steps involved in refining the taxonomic assignment of these MOTUs assigned the 888 MOTUs to 524 different species. Filtering out MOTUs with less than five reads left 411 MOTUS that, when collapsed by their taxonomic assignment, corresponded to 350 different species. Finally, considering only MOTUs with at least five



replicates in a plot and presence in at least two of its replicates as present in that plot left 331 different species.

Rarefaction curves (Figure S1) showed that most plots reached a plateau in all or some of its replicates. The exceptions were the only remaining replicate of PM1 and two of the replicates of PS2, all of which yielded very low numbers of reads.

### 3.2 | Delimitation of adult and juvenile clusters

After collapsing the species that were undistinguishable by the COI fragment, the final number of species for adults and juveniles was 363 and 331, respectively. The combination of both life stages yielded a total of 491 different species. Of those, 160 were found exclusively as adults, 128 exclusively as juveniles and 203 as both life stages. The addition of the species found only as juveniles represented a 35% increase with respect to all the species found as adults. Excluding PM1 (whose sequencing was unsuccessful) the number of matching species varied from 24 (in PM2) to 44 (in PP4) (Figure 2). As for the percentages, PM2 was the plot where juveniles provided the greatest addition of species (169% more species than only with adults), followed by PC2 (162%), PC3 (119%), PC4 (117%), PP3 (112%), PC1 (102%), PP1 (92%), PS2 (83%), PS1 (73%), PO1 (68%), PA1 (58%), PP2 (56%), PO2 (54%), PP4 (47%) and PA2 (44%).

The number of species recovered as juveniles was higher than that of adults in every park (5.4% higher in PO, 8.2% higher in PP, 17.6% higher in PA, 29.8% higher in PS, 45.9% higher in PC and 69.1% higher in PM). Most of the families showed a similar or identical number of species in adults and juveniles, but some specific families showed clear differences (Figure 3). For example, the number of species of juvenile orb-weavers (Araneidae) was much greater than that of adults in all the parks, and a similar trend (although to a lesser extent) was also observed in cob-weaving spiders (Theridiidae) and the sit-and-wait hunting families Philodromidae and Thomisidae crab spiders. Interestingly, for sheet-weaving spiders (Linyphiidae) the number of species represented as adults was greater than that of juveniles in the northern parks (as much as twice as high in PP) but the opposite trend was observed in the southern parks. While the number of sheet-weaving spider MOTUs collected as juveniles remained very similar across all parks, ranging from 12 to 17, the

number of species recovered as adults decreased abruptly from the northern parks (16, 24 and 34) to the southern parks (7, 10 and 11).

### 3.3 | Evaluation of the effect of including juveniles on community patterns

In the adult based NMDS (Figure 4a) the distances between the assemblages of the same park were generally lower than the distances between the parks. This ordination showed a clear separation between northern assemblages and southern assemblages along the first component. The NMDS performed including individuals of all life stages (Figure 4b) did not show a clear north-south distinction, and the parks were less homogeneous than in the NDMS based on adults. The Mantel test revealed that the correlation between the two distance matrices was significantly low ( $r = .336$ ,  $p = .004$ ). ANOSIM analyses revealed significant differences between the species composition of northern and southern assemblages with the data set containing only adult species ( $R = .944$ ,  $p = .0005$ ), but not with the data set including both adults and juveniles ( $R = 0.136$ ,  $p = .078$ ).

### 3.4 | Assessment of species abundance from metabarcoding data

We built beta regression models for the spider families (14) that were present in at least 10 assemblages. Three of the families provided appropriate models for the proportion of both weight and abundance; four families did so only for the models relating proportion of reads and proportion of weight; and three families did so only for the models relating proportion of reads and proportion of individuals (Figure 5), adding to a total of 10 families.

Although the relation between RRA and weight or abundance was positive in all models, its slope varied across spider families. Even in the three families where both models had a  $p > .05$ , the two curves were almost overlapping in two of the families (Lycosidae and Salticidae) but rather different in the third one (Clubionidae). Also, in some families the observations were consistently above or below the 1:1 line. In the plots linking RRA

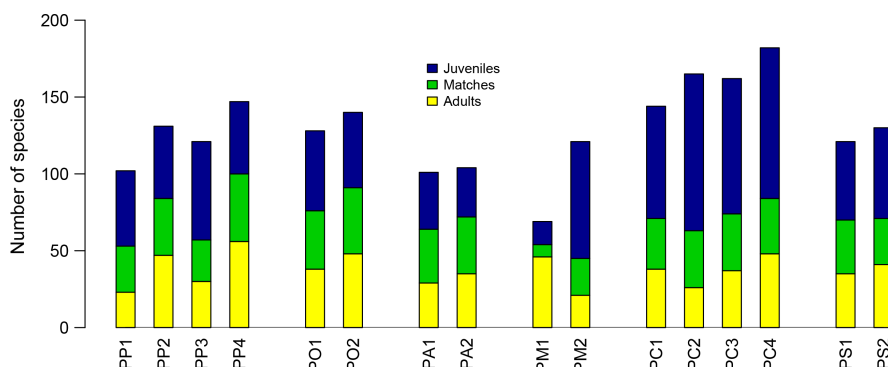
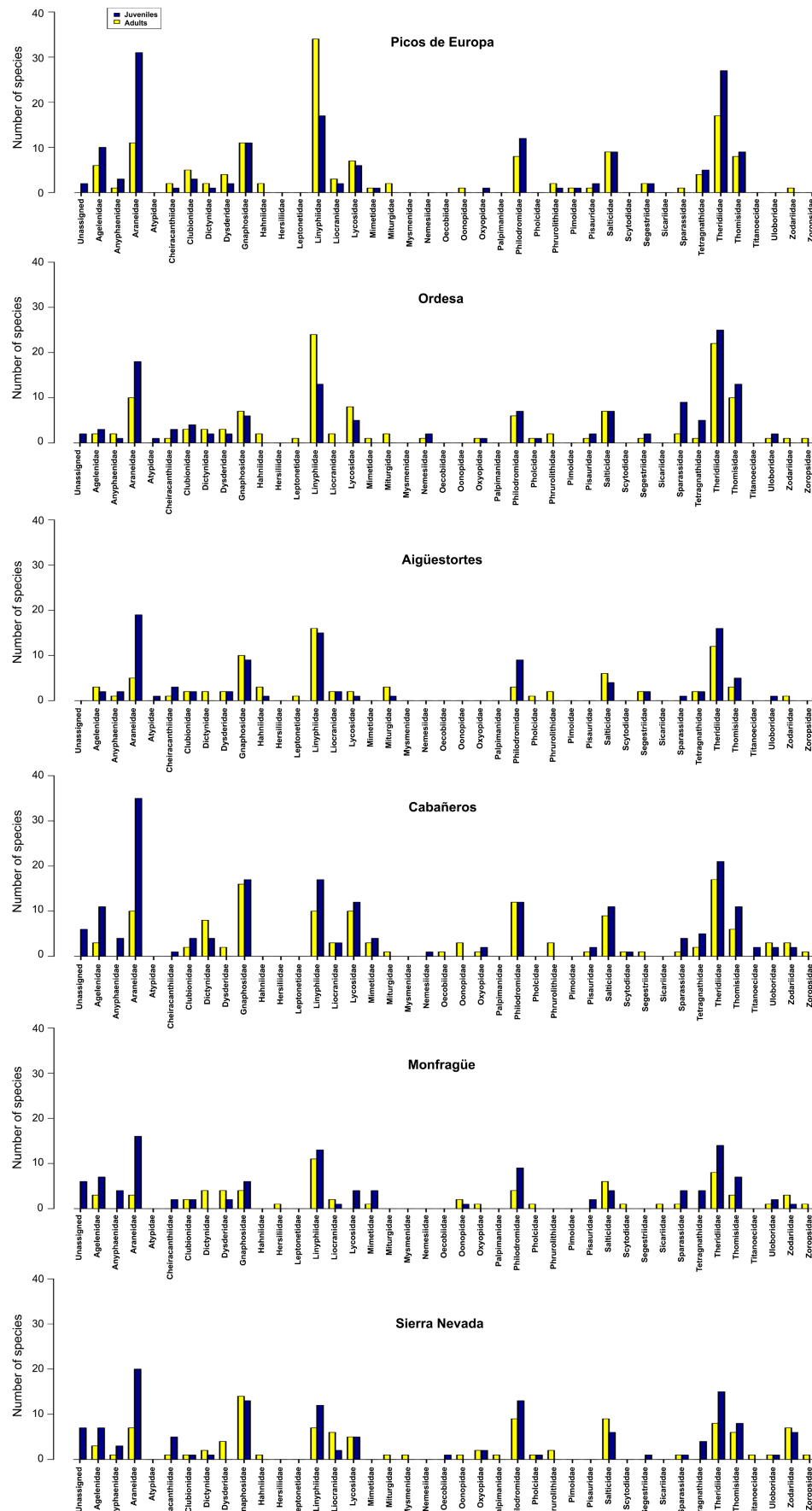
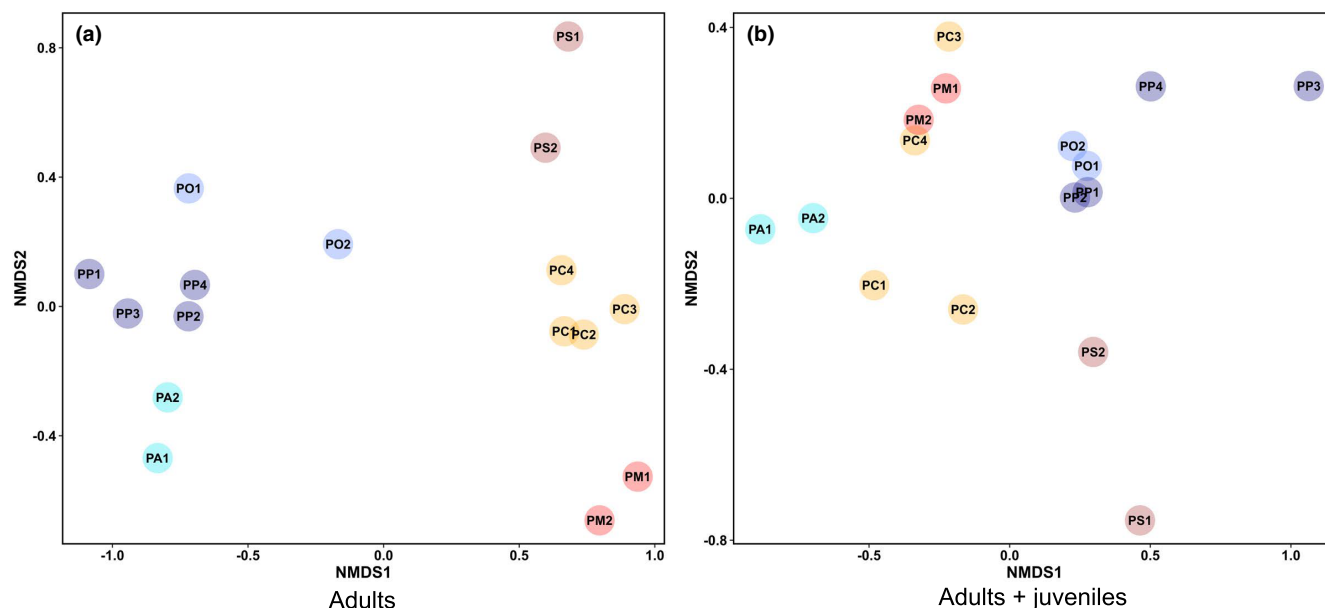


FIGURE 2 Number of species found only as adults, only as juveniles and as both life stages in the 16 sampled plots



**FIGURE 3** Number of species found as adults (yellow) and number of species found as juveniles (blue) for each family in every National Park



**FIGURE 4** Nonmetric multidimensional scaling (NMDS) ordinations of the community composition based on presence/absence data using the Bray-Curtis dissimilarity index. Northern assemblages are represented in cold colours and southern assemblages are represented in warm colours. (a) NMDS using only adult individuals (stress = 0.074). (b) NMDS using all individuals (adults and juveniles) (stress = 0.065)

and abundance of individuals, funnel-web (Agelenidae) and cob-weaving spiders (Theridiidae) were above the 1:1 line, while ghost spiders (Anyphaenidae) and ground-dwellers (Clubionidae) were mostly below. In the plots linking RRA and mass, Clubionidae ground-dwellers, Philodromidae crab spiders and Theridiidae cob-weavers were above the 1:1 line, while Agelenidae funnel-weavers, Araneidae orb-weavers, Gnaphosidae ground-dwellers and Lycosidae wolf spiders were mostly below. The family with the largest juvenile individuals was Agelenidae (18.6 mg per individual), followed by Araneidae (7.6 mg), Lycosidae (6.6 mg), Gnaphosidae (5.5 mg), Salticidae (3 mg), Liocranidae (2.6 mg), Philodromidae (2.4 mg), Theridiidae (2.1 mg), Anyphaenidae (1.5 mg) and Clubionidae (0.9 mg).

## 4 | DISCUSSION

### 4.1 | An important component of biodiversity at a given time is mostly represented by juveniles

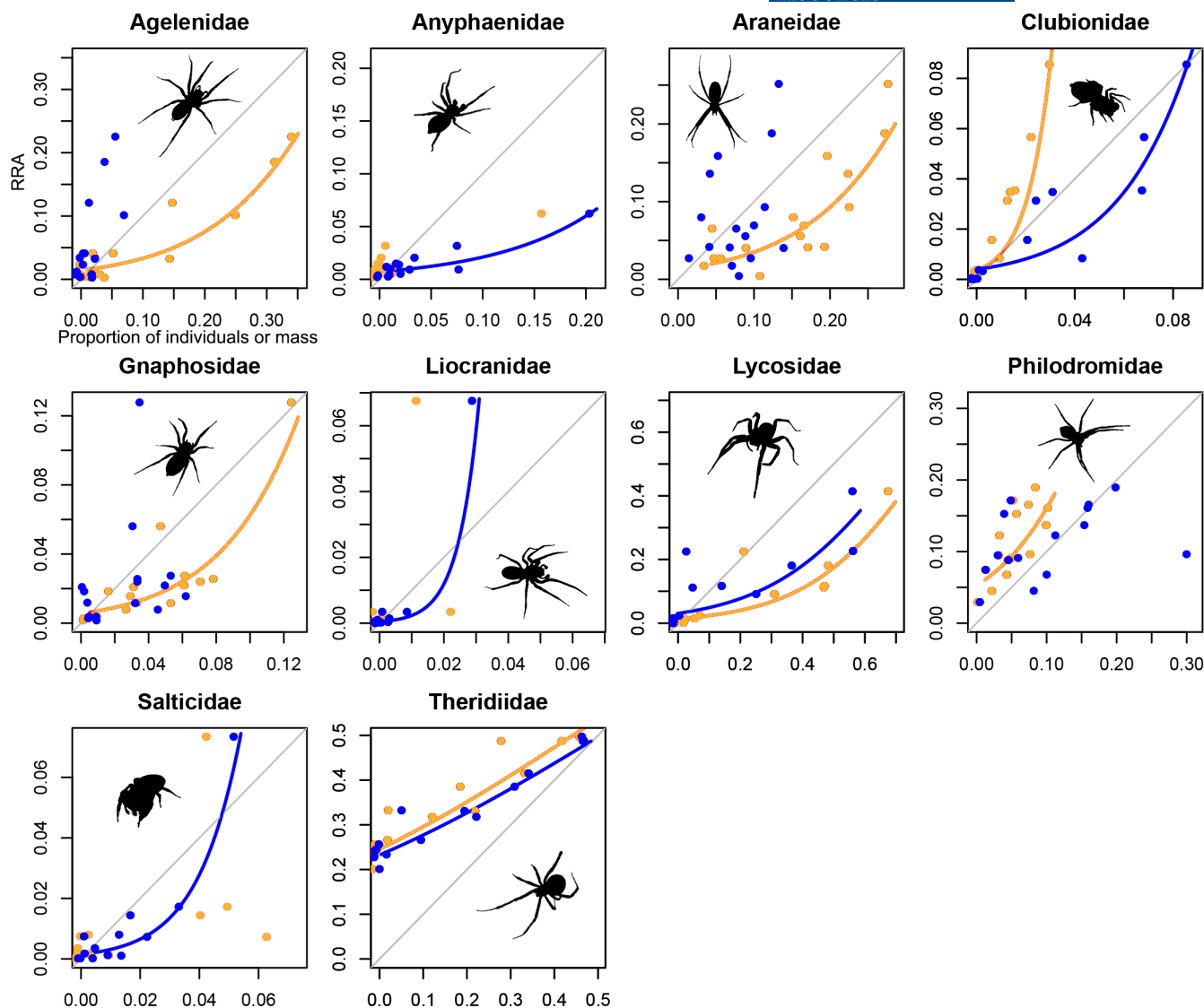
Immature specimens represented 59% of all the captured specimens on average, ranging between 39%–76% per plot. This result is similar to those obtained in other studies (Cardoso et al., 2004; Malumbres-Olarte, Boieiro, et al., 2020; Malumbres-Olarte et al., 2019; Russell-Smith & Stork, 1995; Silva, 1996; Soukainen et al., 2020), which ranged from 40% to 70%, and offers a first insight on the relevance of juvenile stages in spider inventories and the conclusions derived from them.

Overall, the number of species estimated from juveniles by metabarcoding was lower than that of adults. This could suggest that, at the time of the samplings, there were fewer species in the

juvenile stage than in the adult stage. However, this finding could well be an artefact related to the sampling methods, as immature spiders are smaller than adults and may be more difficult to detect by direct sampling techniques. The combination of both life stages yielded a total of 491 different species, 35% higher than the richness we obtained considering only adults in the entire sampling. The degree to which immature stages contributed to species richness, however, was not constant across all our plots. In almost half of the individual samplings, the total number of species including juveniles more than doubled the richness obtained with adults. These results indicate that a large part of the diversity may be ignored by spider bioinventories that are performed exclusively on adults. Interestingly, the contribution of juveniles was higher in southern parks than in northern ones. Although knowing the reasons behind this pattern would require further study, we suspect the difference may be related to the phenological differences between the two latitudes.

Although most families recovered a similar species richness in adult and in juvenile stages, some had important differences (Figure 3). In the case of araneids, the fact that the number of captured juveniles was almost twice as high as the number of adults may have made the number of species captured as immatures greater because it increased the chances of sampling additional species for this stage. The additional diversity found only as juvenile spiders may also be related to the phenology of this family. Larger orb-weavers (araneids) mature in autumn in temperate zones, while smaller species tend to mature earlier (Levi, 1974). This observation fully matches our findings, as most of the araneids that were found exclusively as juveniles are large species of orb-weavers, such as *Argiope lobata* Pallas, 1772, *Larinioides patagiatus* (Clerck, 1757) or several *Araneus* species.





**FIGURE 5** Plots of the relation between relative read abundance (RRA) and proportion of mass (red) and proportion of individuals (blue). Only models with  $p \leq 0.05$  and an adequate goodness of fit are shown

Also interesting is the case of linyphiids sheet-weavers, whose adult species richness was higher than juvenile richness in the northern parks (as much as twice as high in PP), but lower in the southern parks. Linyphiids are known to be much more diverse in temperate regions than in the subtropics and tropics (Cardoso et al., 2011). We indeed found more species in northern parks, which have a more temperate climate, than in southern parks, with a significantly drier and warmer climate. The differences in adult and juvenile richness between north and south, however, may indicate the existence of different predominant phenologies within linyphiids at both latitudes.

Although our data had a large taxonomic scope (order level), there are still some caveats to our analyses. All the MOTUs that could not be assigned to any nominal species could provide important additional information. Nevertheless, it is expected that completeness of reference databases will increase in coming years and the accuracy of taxonomic assignments in metabarcoding studies will improve.

## 4.2 | Including juveniles has an effect on community patterns

The Mantel test revealed that there are substantial differences in the species compositions of assemblages when considering only adult specimens or individuals of all life stages, and the ANOSIM showed that taking immature specimens into account reduced the differences in species composition between assemblages at different latitudes. The NMDS plots also revealed differences in assemblage similarities between the two approaches. While in the NMDS with only adults most assemblages were more similar to other assemblages within the same park than to assemblages from different parks, this pattern was much less pronounced in the NMDS that includes juveniles. In addition, the marked and significant distinction between assemblages at higher latitudes and assemblages at lower latitudes found with adults only, was not as clear when considering all the individuals, although the results were marginally significant.

This less clear North-South split in the NMDS based on specimens of all life stages suggests that phenological differences among communities may exaggerate the differences between assemblage compositions when considering only adult stages. At the time of the sampling, a species might be present as adult in a certain community and still as juvenile in another community, so by considering exclusively adult stages we might be inadvertently accentuating the differences in species composition among assemblages in different regions.

Given that seasonality and marked phenologies with different life stages present at different times of the year are very common across many animal groups (Jakob et al., 2003; Lazaridou-Dimitriadou & Sgardelis, 1997; Scott & Epstein, 1987), we suspect that the phenomenon that we describe here might equally apply to inventories performed on other organisms. Studies on other diverse invertebrate taxa comparing alpha diversity estimates of a community using only individuals of a certain life stage or all individuals might help reveal if this trend is constant across the tree of life.

#### 4.3 | Metabarcoding data may provide abundance information

As expected, and in accordance with other studies investigating the quantitative power of metabarcoding in spiders (Kennedy et al., 2020) or other taxa (Deagle et al., 2019; Krehenwinkel et al., 2017; Lamb et al., 2019; Schenk et al., 2019; Thomas et al., 2016), our results indicate that the relative read abundance (RRA) of a taxon is positively related to its proportion in both weight and abundance. However, the strength of this relation is not constant across spider families, which dissipates the possibility of using a unique correction factor to derive abundance information from read counts for all taxa obtained in metabarcoding analyses. Indeed, similar studies have also found these differences in the factor linking RRA and individual abundance across spider families (Kennedy et al., 2020) or RRA and mass across different taxa in other animal groups (Thomas et al., 2016).

Interestingly, the observations of some of the families in the plots relating RRA to proportion in mass were consistently above or below the 1:1 line. Upon detailed inspection, the spider families mostly above the identity line corresponded to families with small juvenile individuals (between 0.9 and 2.4 mg in mass), while families mostly below the identity line corresponded to those with large juvenile individuals (between 5.5 and 18.7 mg). Families with an intermediate juvenile mass (2.5 to 3 mg) were not clearly above or below the 1:1 line. This suggests that taxa with small or large juvenile sizes might be respectively over- or underrepresented by their reads counts with respect to their real weight proportion in the sample in metabarcoding analyses. Additional studies with spiders and other taxa would help determine if this is a consistent trend and, if so, if it is applicable to other groups apart from spiders.

We adhere to the suggestion made in previous studies that, albeit with caution and with a certain degree of uncertainty, using the

RRA with the corresponding correction factors as a surrogate for the occurrence of a taxon in the sample still provides a more precise information of the community composition than using presence/absence data (Kennedy et al., 2020; Lamb et al., 2019). However, these correction factors need to be developed individually for different taxa, for example using mock communities included as quantitative controls during metabarcoding (Lamb et al., 2019).

## 5 | CONCLUSIONS

Our study suggests that incorporating immature stages of spiders in bioinventorying initiatives has a relevant effect on diversity estimates, because a considerable proportion of the species present as juveniles is not found among adults. This impact goes beyond simply modifying species richnesses, as it also alters the level of similarity among different assemblages regarding species composition. This is probably due to phenological differences associated with different latitudes or elevations, the effect of which is especially true for time-limited samplings. These findings do not question the information provided by adult-based inventories but add a novel and relevant layer of knowledge previously overlooked that may influence some of the interpretations derived from biological inventories. Adding juvenile information to rapid biodiversity assessment protocols provides more accurate data regarding comparisons of community composition. Metabarcoding analysis of all stages present in a sample enables more effective monitoring strategies, and ultimately better-informed conservation decisions.

The proportion of reads obtained from metabarcoding for certain spider families was positively related to their proportion in weight and abundance in the sample, suggesting that metabarcoding data are to a certain extent quantitative. The strength of this relation, however, was not constant across families, as already reported in former studies. Nonetheless, the use of read counts appropriately transformed with taxon-specific correction factors as a proxy of the occurrence of a taxon in the sample could still provide more accurate information about the community composition than simple presence/absence data.

## ACKNOWLEDGEMENTS

Collections were conducted under the corresponding permits kindly provided by the following individuals and institutions: Miguel M. de la Hoz (Picos de Europa), Elena Villagrasa (Ordessa), Maria Merced Aniz Montes (Aigüestortes), Angel Rodriguez Martin (Monfragüe), Angel Gómez Manzaneque (Cabañeros), and Blanca Ramos Losada (Sierra Nevada). We are grateful to all the people that contributed to the samplings as well as to all the park rangers that supported us in the field.

This work was funded by the Organismo Autónomo de Parques Nacionales (OAPN no. 485/2012) (to MA). MD was supported by APIF PhD fellowship from the University of Barcelona. Additional support was provided by 2017SGR73 (to MA) from the Catalan Government.

## AUTHOR CONTRIBUTIONS

Miquel A. Arnedo, Jagoba Malumbres-Olarte and Marc Domènech designed the study. Alba Enguñados and Marc Domènech performed the laboratory work. Owen S. Wangensteen, Marc Domènech and Jagoba Malumbres-Olarte performed the data analyses. Marc Domènech, Jagoba Malumbres-Olarte and Miquel A. Arnedo drafted the manuscript with contributions from all authors. All authors revised and approved the final manuscript.

## CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## OPEN RESEARCH BADGES



This article has earned an Open Data, for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at <https://datadryad.org/stash/share/MIRsXuxmuddQXSJay1ywfjmQOniHTjHaiSfir1a4Ny4>.

## DATA AVAILABILITY STATEMENT

Raw fastq files for all samples are available from the SRA archive in Genbank (Accession: [PRJNA826515](https://www.ncbi.nlm.nih.gov/PRJNA826515)). Fasta file with sequences of adult specimens: Dryad, <https://datadryad.org/stash/share/MIRsXuxmuddQXSJay1ywfjmQOniHTjHaiSfir1a4Ny4>. Fasta file with sequences of immature specimens: Dryad, <https://datadryad.org/stash/share/MIRsXuxmuddQXSJay1ywfjmQOniHTjHaiSfir1a4Ny4>.

## ORCID

Marc Domènech  <https://orcid.org/0000-0003-4352-5301>

Miquel A. Arnedo  <https://orcid.org/0000-0003-1402-4727>

## REFERENCES

- Aitchison, C. W. (1984). The phenology of winter active spiders. *Journal of Arachnology*, 12, 249–271.
- Antich, A., Palacín, C., Wangensteen, O. S., & Turon, X. (2021). To denoise or to cluster? That is not the question. Optimizing pipelines for COI metabarcoding and metaphylogeography. *BMC Bioinformatics*, 2021, 22–177. <https://doi.org/10.1186/s12859-021-04115-6>
- Bakker, J., Wangensteen, O. S., Baillie, C., Buddo, D., Chapman, D. D., Gallagher, A. J., Guttridge, T. L., Hertler, H., & Mariani, S. (2019). Biodiversity assessment of tropical shelf eukaryotic communities via pelagic eDNA metabarcoding. *Ecology and Evolution*, 9, 14341–14355. <https://doi.org/10.1002/ece3.5871>
- Barlow, J., Lennox, G. D., Ferreira, J., Berenguer, E., Lees, A. C., Mac Nally, R., Thomson, J. R., Ferraz, S. F. D. B., Louzada, J., Oliveira, V. H. F., Parry, L., De Castro, R., Solar, R., Vieira, I. C. G., Aragañ, L. E. O. C., Begotti, R. A., Braga, R. F., Cardoso, T. M., de Oliveira Jr, R. C., ... Gardner, T. A. (2016). Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature*, 535, 144–147. <https://doi.org/10.1038/nature18326>
- Barnosky, A. D., Hadly, E. A., Bascompte, J., Berlow, E. L., Brown, J. H., Fortelius, M., Getz, W. M., Harte, J., Hastings, A., Marquet, P. A., Martinez, N. D., Mooers, A. O., Roopnarine, P. D., Vermeij, G., Williams, J. W., Gillespie, R. G., Kitzes, J., Marshall, C., Matzke, N., ... Smith, A. B. (2012). Approaching a state shift in Earth's biosphere. *Nature*, 486, 52–58. <https://doi.org/10.1038/nature11018>
- Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, 471, 51–57. <https://doi.org/10.1038/nature09678>
- Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W., & de Bruyn, M. (2014). Environmental DNA for wild-life biology and biodiversity monitoring. *Trends in Ecology and Evolution*, 29, 358–367. <https://doi.org/10.1016/j.tree.2014.04.003>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16, 176–182.
- Cardoso, P. (2009). Standardization and optimization of arthropod inventories—the case of Iberian spiders. *Biodiversity and Conservation*, 18, 3949–3962. <https://doi.org/10.1007/s10531-009-9690-7>
- Cardoso, P., Borges, P. A. V., Triantis, K. A., Ferrández, M. A., & Martín, J. L. (2012). The underrepresentation and misrepresentation of invertebrates in the IUCN Red List. *Biological Conservation*, 149, 147–148. <https://doi.org/10.1016/j.biocon.2012.02.011>
- Cardoso, P., Pekár, S., Jocqué, R., & Coddington, J. A. (2011). Global patterns of guild composition and functional diversity of spiders. *PLoS One*, 6, e21710.
- Cardoso, P., Silva, I., De Oliveira, N. G., & Serrano, A. R. M. (2004). Indicator taxa of spider (Araneae) diversity and their efficiency in conservation. *Biological Conservation*, 120, 517–524. <https://doi.org/10.1016/j.biocon.2004.03.024>
- Chapin III, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C., & Díaz, S. (2000). Consequences of changing biodiversity. *Nature*, 405, 234–242. <https://doi.org/10.1038/35012241>
- Clerck, C. (1757). Aranei Svecici. Svenska spindlar, uti sina hufvud-slagter indelte samt under några och sextio särskildte arter beskrefne och med illuminerade figurer uplyste. *Laurentius Salvius, Stockholmiae [= Stockholm]*, 154. <https://doi.org/10.5962/bhl.title.119890>
- Coddington, J. A., Young, L. H., & Coyle, F. A. (1996). Estimating spider species richness in a southern Appalachian cove hardwood forest. *Journal of Arachnology*, 24, 111–128.
- Crespo, L., Domènech, M., Enguñados, A., Malumbres-Olarte, J., Cardoso, P., Moya-Laraño, J., Frías-López, C., Macías-Hernández, N., De Mas, E., Mazzuca, P., Mora, E., Opatova, V., Planas, E., Ribera, C., Roca-Cusachs, M., Ruiz, D., Sousa, P., Tonzo, V., & Arnedo, M. (2018). A DNA barcode-assisted annotated checklist of the spider (Arachnida, Araneae) communities associated to white oak woodlands in Spanish National Parks. *Biodiversity Data Journal*, 1–273. <https://doi.org/10.3897/BDJ.6.e29443>
- Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Handbook of Meta-Analysis*, 34, 129–150.
- Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzin, T. R., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, 28, 391–406.
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, 345, 401–406. <https://doi.org/10.1126/science.1251817>
- Dobyns, J. R. (1997). Effects of sampling intensity on the collection of spider (Araneae) species and the estimation of species richness. *Environmental Entomology*, 26, 150–162. <https://doi.org/10.1093/ee/26.2.150>
- Eberhard, W. G. (1985). *Sexual selection and animal genitalia*. Harvard University Press.
- Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and

- biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS One*, 10, 1–16. <https://doi.org/10.1371/journal.pone.0130324>
- Frøsvlev, T. G., Kjølter, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, 8(1188), 1–11.
- Garcés-Pastor, S., Wangenstein, O. S., Pérez-Haase, A., Pélachs, A., Pérez-Obiol, R., Cañellas-Boltà, N., Mariani, S., & Vegas-Vilarrúbia, T. (2019). DNA metabarcoding reveals modern and past eukaryotic communities in a high-mountain peat bog system. *Journal of Paleolimnology*, 62, 425–441. <https://doi.org/10.1007/s10933-019-00097-x>
- Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, 13, 851–861.
- Hebert, P., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, 270, 96–99.
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518–522. <https://doi.org/10.1093/molbev/msx281>
- Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, 46, 10–18. <https://doi.org/10.1111/j.1365-2664.2008.01600.x>
- Jakob, C., Poizat, G., Veith, M., Seitz, A., & Crivelli, A. J. (2003). Breeding phenology and larval distribution of amphibians in a Mediterranean pond network with unpredictable hydrology. *Hydrobiologia*, 499, 51–61.
- Kennedy, S., Prost, S., Overcast, I., Rominger, A. J., Gillespie, R. G., & Krehenwinkel, H. (2020). High-throughput sequencing for community analysis: the promise of DNA barcoding to uncover diversity, relatedness, abundances and interactions in spider communities. *Development Genes and Evolution*, 230, 185–201. <https://doi.org/10.1007/s00427-020-00652-x>
- Kemp, J., López-Baucells, A., Rocha, R., Wangenstein, O. S., Andriatafika, Z., Nair, A., & Cabeza, M. (2019). Bats as potential suppressors of multiple agricultural pests: a case study from Madagascar. *Agriculture, Ecosystems & Environment*, 269, 88–96.
- Koch, C. L. (1836). *Die Arachniden* (pp. 1–104). Nürnberg: Dritter Band. C. H. Zeh'sche Buchhandlung.
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7, 1–12. <https://doi.org/10.1038/s41598-017-17333-x>
- Kremen, C., Colwell, R. K., Erwin, T. L., Murphy, D. D., Noss, R. F., & Sanjayan, M. A. (1993). Terrestrial arthropod assemblages: Their use in conservation planning. *Conservation Biology*, 7, 796–808. <https://doi.org/10.1046/j.1523-1739.1993.740796.x>
- Kulikova, T., Aldebert, P., Althorpe, N., Baker, W., Bates, K., Browne, P., van den Broek, A., Cochrane, G., Duggan, K., Eberhardt, R., Faruque, N., Garcia-Pastor, M., Harte, N., Kanz, C., Leinonen, R., Lin, Q., Lombard, V., Lopez, R., Mancuso, R., ... Apweiler, R. (2004). The EMBL nucleotide sequence database. *Nucleic Acids Research*, 32, 27–30. <https://doi.org/10.1093/nar/gkh120>
- Kuntner, M., & Baxter, I. H. (1997). A preliminary investigation of spider species richness in an eastern Slovenian broadleaf forest. *Proceedings of the 16th European Colloquium of Arachnology*, 173–182.
- Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology*, 28, 420–430. <https://doi.org/10.1111/mec.14920>
- Lazaridou, M., & Sgardelis, S. (1997). Phenological patterns and life history tactics of Helicoidea (Gastropoda, Pulmonata) snails from Northern Greece. *Iberus*, 15(2), 25–34.
- Leather, S. R. (2018). "Ecological Armageddon" – more evidence for the drastic decline in insect numbers. *Annals of Applied Biology*, 172, 1–3. <https://doi.org/10.1111/aab.12410>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10, 1–14. <https://doi.org/10.1186/1742-9994-10-34>
- Levi, H. W. (1974). The orb-weaver genera Araniella and Nuctenea. *Bulletin of the Museum of Comparative Zoology*, 146, 291–316.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2015). Swarmv2: Highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3, e1420. <https://doi.org/10.7717/peerj.1420>
- Malumbres-Olarte, J., Boieiro, M., Cardoso, P., Carvalho, R., Crespo, L. C., Gabriel, R., Hernández, N. M., Paulo, O. S., Pereira, F., Rego, C., Ros-Prieto, A., Silva, I., Vieira, A., Rigal, F., & Borges, P. A. V. (2020). Standardised inventories of spiders (Arachnida, Araneae) of Macaronesia II: The native forests and dry habitats of Madeira archipelago (Madeira and Porto Santo islands). *Biodiversity Data Journal*, 8, 3973–3996. <https://doi.org/10.3897/BDJ.8.e47502>
- Malumbres-Olarte, J., Cardoso, P., & Crespo, L. C. (2019). Standardised inventories of spiders (Arachnida, Araneae) of Macaronesia I: The native forests of the Azores (Pico and Terceira islands). *Biodiversity Data Journal*, 7, e32625.
- Malumbres-Olarte, J., Crespo, L. C., Domènech, M., Cardoso, P., Moya-Laraño, J., Ribera, C., & Arnedo, M. A. (2020). How Iberian are we? Mediterranean climate determines structure and endemism of spider communities in Iberian oak forests. *Biodiversity and Conservation*, 29, 3973–3996. <https://doi.org/10.1007/s10531-020-02058-7>
- Meiklejohn, K. A., Wallman, J. F., & Dowton, M. (2012). DNA barcoding identifies all immature life stages of a forensically important flesh fly (Diptera: Sarcophagidae). *Journal of Forensic Sciences*, 58, 184–187. <https://doi.org/10.1111/j.1556-4029.2012.02220.x>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*, New Orleans, 2010, 1–8.
- Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 305, 1188–1195. <https://doi.org/10.1093/molbev/mst024>
- Morinière, J., Cancian De Araujo, B., Lam, A. W., Hausmann, A., Balke, M., Schmidt, S., Hendrich, L., Doczkal, D., Fartmann, B., Arvidsson, S., & Haszprunar, G. (2016). Species identification in malaise trap samples by DNA barcoding based on NGS technologies and a scoring matrix. *PLoS One*, 11, 1–14. <https://doi.org/10.1371/journal.pone.0155497>
- Nadal, M. F., Achitte-Schmutzler, H. C., Zanone, I., Gonzalez, P. Y., & Avalos, G. (2018). Diversidad estacional de arañas en una reserva natural del Espinal en Corrientes, Argentina. *Caldasia*, 40, 129–143. <https://doi.org/10.15446/caldasia.v40n1.67362>
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274. <https://doi.org/10.1093/molbev/msu300>
- Norris, K. C. (1999). Quantifying change through time in spider assemblages: Sampling methods, indices and sources of error. *Journal of Insect Conservation*, 3, 309–325.



- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G. L., & Wagner, H. (2010). Vegan: community ecology package. R package version.
- Oliver, I., & Beattie, A. J. (1996). Designing a cost-effective invertebrate survey: A test of methods for rapid assessment of biodiversity. *Ecological Applications*, 6, 594–607. <https://doi.org/10.2307/2269394>
- Pallas, P. S. (1772). *Spicilegium zoologica. Tomus 1. Continens quadrupedum, avium, amphibiorum, piscium, insectorum, molluscorum aliorumque marinorum fasciculos decem. Berolini*, 1(9), 44–50.
- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12), 1373–1379. [https://doi.org/10.1016/s0895-4356\(96\)00236-3](https://doi.org/10.1016/s0895-4356(96)00236-3)
- Pétillon, J., Leroy, B., Djoudi, E. A., & Vedel, V. (2018). Small and large spatial scale coexistence of ctenid spiders in a neotropical forest (French Guiana). *Tropical Zoology*, 31(2), 85–98. <https://doi.org/10.1080/03946975.2018.1448531>
- Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15, 819–830. <https://doi.org/10.1111/1755-0998.12355>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rádai, Z., Kiss, B., & Samu, F. (2017). Effect of weather conditions on cohort splitting in a Wolf spider species. *Journal of Arachnology*, 45, 444–447. <https://doi.org/10.1636/JoA-S-17-008.1>
- Ratcliffe, F. C., Uren Webster, T. M., Rodriguez-Barreto, D., O'Rourke, R., Garcia de Leaniz, C., & Consuegra, S. (2020). Rapid quantitative assessment of fish larvae community composition using metabarcoding. *bioRxiv*. <https://doi.org/10.1101/2020.01.01.884544>
- Ratnasingham, S., & Hebert, P. (2007). BOLD: The barcode of life data system (<http://www.boldsystems.org>). *Molecular Ecology Notes*, 7, 355–364.
- Richard, B., Decaëns, T., Rougerie, R., James, S. W., Porco, D., & Hebert, P. (2010). Re-integrating earthworm juveniles into soil biodiversity studies: Species identification through DNA barcoding. *Molecular Ecology Resources*, 10, 606–614. <https://doi.org/10.1111/j.1755-0998.2009.02822.x>
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. <https://doi.org/10.7717/peerj.2584>
- Russell-Smith, A., & Stork, N. E. (1995). Composition of spider communities in the canopies of rainforest trees in Borneo. *Journal of Tropical Ecology*, 11, 223–235. <https://doi.org/10.1017/S0266467400008695>
- Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., Wangenstein, O. S., Baillie, C., Bryce, E., Dawson, D. A., Ochu, E., Hänfling, B., Lawson Handley, L., Mariani, S., Lambin, X., Sutherland, C., & McDevitt, A. D. (2020). Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *Journal of Applied Ecology*, 57, 707–716. <https://doi.org/10.1111/1365-2664.13592>
- Schaefer, M. (1976). Experimentelle Untersuchungen zum Jahreszyklus und zur Überwinterung von Spinnen (Anarenida). *Zoologische Jahrbücher, Abteilung Systematik, Ökologie Und Geographie Der Tiere*, 103, 127–289.
- Schenk, J., Geisen, S., Kleinbölting, N., & Traunspurger, W. (2019). Metabarcoding data allow for reliable biomass estimates in the most abundant animals on earth. *Metabarcoding and Metagenomics*, 3, 117–126. <https://doi.org/10.3897/mbmg.3.46704>
- Schiebelhut, L. M., Abboud, S. S., Gómez-Daglio, L. E., Swift, H. F., & Dawson, M. N. (2017). A comparison of DNA extraction methods for high-throughput DNA analyses. *Molecular Ecology Resources*, 17, 721–729. <https://doi.org/10.1111/1755-0998.12620>
- Scott, J. A., & Epstein, M. E. (1987). Factors affecting phenology in a temperate insect community. *American Midland Naturalist*, 117, 103–118. <https://doi.org/10.2307/2425712>
- Siegenthaler, A., Wangenstein, O. S., Benvenuto, C., Campos, J., & Mariani, S. (2019). DNA metabarcoding unveils multiscale trophic variation in a widespread coastal opportunist. *Molecular Ecology*, 28, 232–249.
- Silva, D. (1996). Species composition and community structure of Peruvian rainforest spiders: A case study from a seasonally inundated forest along the Samiria river. *Revue Suisse De Zoologie*, 2, 597–610.
- Socolar, J. B., Valderrama Sandoval, E. H., & Wilcove, D. S. (2019). Overlooked biodiversity loss in tropical smallholder agriculture. *Conservation Biology*, 33, 1338–1349. <https://doi.org/10.1111/cobi.13344>
- Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H. K., Epp, L., Schröder-Nielsen, A., Brochmann, C., & Yoccoz, N. G. (2015). Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS One*, 10, 1–18. <https://doi.org/10.1371/journal.pone.0115335>
- Sørensen, L. L., Coddington, J. A., & Scharff, N. (2002). Inventorying and estimating subcanopy spider diversity using semiquantitative sampling methods in an afro-montane forest. *Environmental Entomology*, 31, 319–330. <https://doi.org/10.1603/0046-225X.31.2.319>
- Soukainen, A., Pajunen, T., Korhonen, T., Saarinen, J., Chichorro, F., Jalonen, S., Kiljunen, N., Koskivirta, N., Kuurne, J., Leinonen, S., Salonen, T., Yrjölä, V., Fukushima, C., & Cardoso, P. (2020). Standardised spider (Arachnida, Araneae) inventory of Lammi, Finland. *Biodiversity Data Journal*, 8, 1–15. <https://doi.org/10.3897/BDJ.8.e50775>
- Taylor, R. J., & Doran, N. (2001). Use of terrestrial invertebrates as indicators of the ecological sustainability of forest management under the Montreal Process. *Journal of Insect Conservation*, 5, 221–231.
- Telfer, A., Young, M., Quinn, J., Perez, K., Sobel, C., Sones, J., Levesque-Beaudin, V., Derbyshire, R., Fernandez-Triana, J., Rougerie, R., Thevanayagam, A., Boskovic, A., Borisenko, A., Cadel, A., Brown, A., Pages, A., Castillo, A., Nicolai, A., Glenn Mockford, B. M., ... de-Waard, J. (2015). Biodiversity inventories in high gear: DNA barcoding facilitates a rapid biotic survey of a temperate nature reserve. *Biodiversity Data Journal*, 3, e6313. <https://doi.org/10.3897/BDJ.3.e6313>
- Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2016). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources*, 16, 714–726. <https://doi.org/10.1111/1755-0998.12490>
- Toti, D. S., Coyle, F. A., & Miller, J. A. (2000). A structured inventory of apalachian grass bald and heath bald spider assemblages and a test of species richness estimator performance. *Journal of Arachnology*, 28, 239–345.
- Wangenstein, O. S., Palacín, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hard-bottom communities: High diversity and database gaps revealed by two molecular markers. *PeerJ*, 6, e4705. <https://doi.org/10.7717/peerj.4705>
- Watts, C., Dopheide, A., Holdaway, R., Davis, C., Wood, J., Thornburrow, D., & Dickie, I. A. (2019). DNA metabarcoding as a tool for invertebrate community monitoring: A case study comparison with conventional techniques. *Austral Entomology*, 58, 675–686. <https://doi.org/10.1111/aen.12384>
- Yellareddygar, S. K., Pasche, J. S., Taylor, R. J., Hua, S., & Gudmestad, N. C. (2016). Beta regression model for predicting the development of



- pink rot in potato tubers during storage. *Plant Disease*, 100, 1118–1124. <https://doi.org/10.1094/PDIS-06-15-0696-RE>
- Yen, A., & Butcher, R. (1997). *An overview of the conservation of Non-marine Invertebrates in Australia*. Canberra, Australia, Environment Australia.
- Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3, 613–623. <https://doi.org/10.1111/j.2041-210X.2012.00198.x>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Domènech, M., Wangenstein, O. S., Enguídanos, A., Malumbres-Olarte, J., & Arnedo, M. A. (2022). For all audiences: Incorporating immature stages into standardised spider inventories has a major impact on the assessment of biodiversity patterns. *Molecular Ecology Resources*, 22, 2319–2332. <https://doi.org/10.1111/1755-0998.13625>